

2020年度 助成金交付について

当財団の助成事業を、今年も実施いたします。当財団では、助成総件数690件に達し、助成総金額は5億7,545万円となりました。学術、ひいては社会の発展のためにお役立てくだされば幸いです。ご応募を心よりお待ちしております。

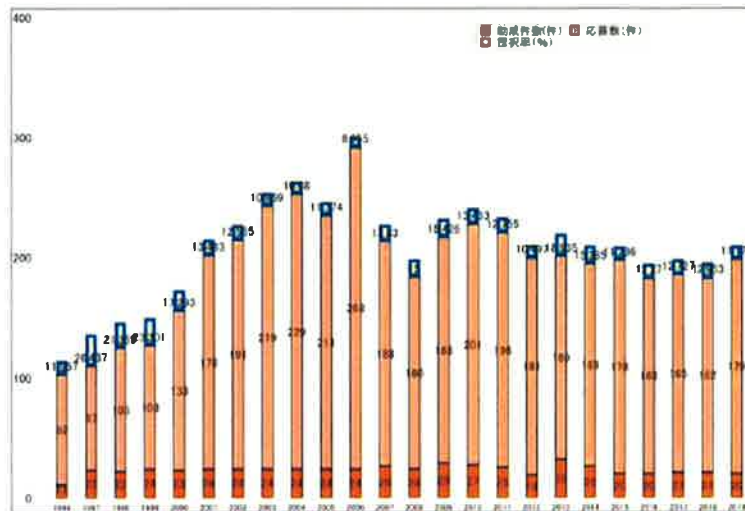
詳細は財団のホームページをご覧ください。

【申請書受付期間:2020年6月1日(月)~8月31日(月)】

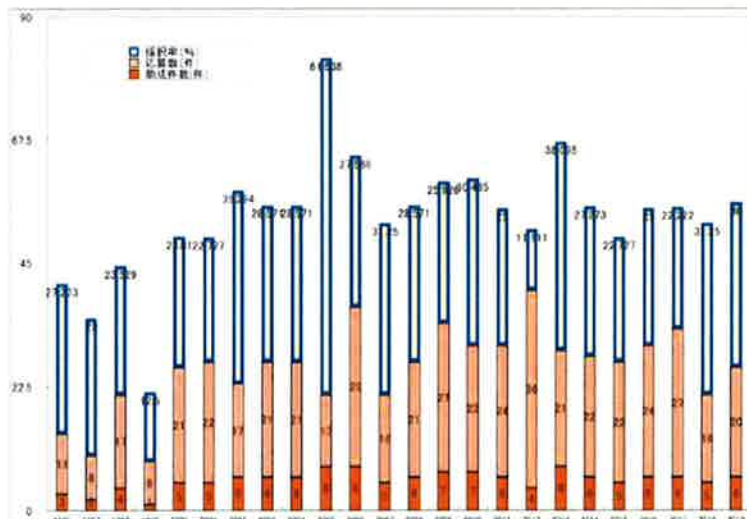
応募手続き

- ◆財団所定の書式(当財団のHPに掲載)を用いて、必要事項を記入して財団事務局あてにEメールで提出してください(郵送も可)。
- ◆申請書の受付完了報告をEメールにてお知らせいたします。
- ◆申請書受付期間外に到着した申請書につきましては受付できませんのでご注意ください。また、電話などで申請書到着遅延、期間外の受付などのお問い合わせはお断りいたします。

■研究助成の推移



■フォーラム・シンポジウム等開催助成の推移



応募要領

〈1〉研究助成

◆応募の資格

助成の対象となる研究を、計画に従って遂行する能力のある方(研究グループを含む)。

◆テーマ・内容

◎情報科学に関する調査、研究および開発で、学術的發展に寄与するものであること。

◎研究の計画および方法が、当該研究の目的を達成するために適切であり、かつ十分な成果が期待できるものであること。

◆助成金の額

助成総額2000万円までを原則とし、選考結果に基づき助成額を決定します。

◆選考結果通知

2020年11月中旬の予定。全員にお知らせします。

◆対象となる経費

機械器具装置の購入費および賃貸料、旅費、消耗品費、謝金等。

◆研究完了日

助成金の交付決定後2年以内。

◆研究成果の帰属

助成研究によって取得された知的財産権は、研究実施者に帰属することとします。ただし、助成研究成果を特許、実用新案または意匠登録として出願し、その後、特許権、実用新案権または意匠権を取得したときは、速やかにその旨を当財団に届け出てください。また、当財団では、「特許庁長官指定学術団体」として指定されていますので、当財団が主催または共催する研究集会で文書をもって発表した場合、発表後6ヶ月以内に特許、実用新案の出願をされたときは、その発明または考案は新規性の喪失の例外とされています。その場合、当財団の証明書が必要となりますのでお申し出ください。

◆その他、留意していただく事項

①研究の成功・不成功にかかわらず助成金の返還は求めませんが、当該研究が実施されなかったり、研究実施者が当財団の規程等に違反した場合には、助成金の一部または全額を返還していただくことがあります。

②助成研究完了の日から起算して30日以内に、完了報告書の提出をお願いします。

③研究の成果を当財団の機関誌等に記載したり、講演会等で発表していただくことがあります。

④助成研究の成果を学会等で発表したり論文にまとめたりする場合は、財団の助成を受けて遂行されたことを明示してください。

⑤応募者の機会均等化を期するため、採択された方は、原則としてその年度後3年間は、選考の対象とされません。

〈2〉フォーラム・シンポジウム等開催助成

◆応募の条件

情報科学に関する学術的發展に寄与するフォーラム・シンポジウム等で、2020年12月1日から2022年3月末日までに開催されること。

◆助成金の額

年度内助成総額200万円までを原則とし、選考結果に基づき、助成額を決定します。

◆選考結果通知

2020年11月上旬の予定。

◆対象となる経費

謝金、旅費、会場費、人件費、消耗品費、印刷製本費、通信運搬費等。

◆その他、留意していただく事項

①フォーラム・シンポジウム等の終了後3ヶ月以内に報告書を提出してください。

②フォーラム・シンポジウム等開催の資料は、申請時に添付のほか、印刷物を発行する場合は送付してください。

③応募者の機会均等化を期するため、採択された団体等は、原則としてその年度後3年間は、選考の対象とされません。

ロボットシンポジウム2020名古屋

ロボット・AIが未来の社会を拓く

開催日 2020年2月5日(水)～2020年2月6日(木)

場所 吹上ホール(名古屋市中小企業振興会館)



開催趣旨

ヒューマンロボットコンソーシアム 会長
ロボットシンポジウム2020名古屋実行委員会 委員長 稲垣 康善

AI、ロボット、IoTなど情報技術の急速な進化で、仲間にしてきたAI・ロボットがまさに仲間となる日常が目前に迫ってきたように思われます。AI・ロボット技術は、第4次産業革命実現のための基盤技術の要と目され、グローバル化、人口減少、少子高齢化などによって地域経済社会が抱える課題の解決、産業の国際競争力の強化、そして次世代産業創生のキー・テクノロジーとして、業種・業態を超えて多くの企業において、研究開発、実用化が進められています。また政府も、成長戦略の柱として、産業や社会生活への利活用促進、さらには産業化のための研究開発を、省庁連携、産学連携などにより積極的に推進しています。

こうした中で、人工知能・ロボット関連の最先端・最新技術の情報交換や技術交流、併せて、人工知能により進化を続けているロボットの実演展示等を行うことにより、人に寄り添い、人に優しく、人と共生する次世代の人工知能・ロボットの研究開発の振興と産業における利活用の促進を図ることを目的として、昨年に引き続き本年もロボットシンポジウムを開催いたします。

本シンポジウムの開催にあたり、格別のご支援、ご協力を賜りました関係者の方々に厚く御礼申し上げますとともに、シンポジウムにご参加の皆様にとりまして有意義で実り多きシンポジウムとなりますことを祈念いたします。

会 議 開 催 報 告

「第20回 理事会」開催

新型コロナウイルスの感染拡大が懸念され始めたことから、集合型会議は自粛して、第20回理事会を決議の省略により実施いたしました。

理事長より各理事監事に対して、

- ① 2020年度事業計画書案及び収支予算書案、資金調達及び設備投資の見込みの承認の件
- ② 第20回評議員会の日時及び場所並びに目的である事項決定の件

が書面にて提案され、全理事からの同意書と全監事から異議のない旨回答書を得て、2020年2月19日に決議の省略が成立しました。

「第20回 評議員会」開催

日本政府から緊急事態宣言も発出されたこともあり、新型コロナウイルスの感染リスクを考慮して、第20回評議員会を決議の省略により実施いたしました。

理事長より評議員に対して、

- ① 2020年度事業計画書案及び収支予算書案、資金調達及び設備投資の見込みの承認の件
- が書面にて提案され、全評議員からの同意書を得て、2020年3月5日に決議の省略が成立しました。

「第7回 臨時理事会」開催

2020年3月18日(水)午前10より、WEB会議システムで第7回臨時理事会を開催いたしました。

代表理事による自己の職務の執行状況報告として、理事長から2019年度の公益目的事業の執行内容を、専務理事から2019年度の財団会計管理の執行内容をそれぞれ報告いたしました。

フォーラム・シンポジウム等開催助成完了報告

(いずれも提出原文のまま、所属は提出時のもの)

7th International Conference on Human-Agent Interaction K30FSXXⅢ第122号

開催責任者：神田 智子(大阪工業大学 教授)
開催期間：2019年10月6日～2019年10月10日
会場と所在地：京都工芸繊維大学3号館(10月6日 ワークショップ)
〒606-8585 京都市左京区松ヶ崎橋上町
京都産業大学むすびわざ館(10月6日～10日 本会議)
〒600-8533 京都府京都市下京区中堂寺命婦町1-10
参加人員：本会議 141名、ワークショップ 73名

成果：

近年、人とインタラクションを行うロボットやソフトウェアエージェントが日常生活に入り込んできている。しかし、そのインタラクションデザイン(設計原理)についてはいまだ十分な研究が行われていない。本会議は、人とロボットやエージェントとのインタラクションや、電子メディアを介した人と人のコミュニケーションに関する、分野横断的な研究を推進することを目的に開催された。

本会議のプログラムは、以下に記載されているように、研究発表13セッション(口頭発表8セッション、ポスター発表3セッション)および招待講演2件から構成された。以下に、研究発表の口頭発表セッションを中心に会議の概要について報告する。

まず、HAIにおける対話デザイン、インタラクションデザインに関するセッションが3つ設けられた。具体的なセッション名は、Learning Interaction, Conversational Design, Learning & Conversationである。HAI研究の基盤となるのは、言うまでもなく、円滑な人とエージェントの対話インタラクションの探求である。この意味でも、全10セッションのうち3セッションがこれらのテーマであったことは有意義であったと考える。

次に、ロボットやエージェントと人間との共感や感情、およびそれらの知覚や認知に関するセッションが3つ設けられた。具体的なセッション名としては、Emotion & Empathy 1、2、Perceptionである。HAI研究

が、人間とエージェントやロボットとの対話から、感情の表出や共感行動にいたるまでより深いインタラクションを目指して進化していることを表しているといえる。さらに、エージェントの動作、行動設計やそれらのコンピューターモデルに関するセッションが2つ設けられた。具体的なセッション名としては、Computational Model, Movements & Behaviorsである。人とインタラクションを行うロボットやエージェントを設計するうえで、技術的なブレークスルーが当然必要となる。これらのセッションでは、この目的を目指した技術的な研究成果発表が行われた。

最後に、本会議では、国際会議としては初めての試みである、機械翻訳システムを用いた質疑応答支援を行った。参加者が母国語で口頭で質問をすると、リアルタイムで翻訳結果が日英翻訳されスクリーンに表示されるというものである。本会議では、MT Sessionというセッション名で7つの発表で支援が行われた。

さらに、HAIに関連した研究分野のエキスパートによる2件の招待講演が、本会議の議論をさらに活性化し、有意義なものとしていた。

上記の会議内容でも述べたとおり、貴財団からの助成のおかげで、本会議は当初の目的を達成し、さらに本実行委員会が予想した以上の参加者、投稿論文数があり、5日間にわたり有意義な議論が交わされていた。今後はさらに海外の研究者の参加を促し、議論を活発なものにしていきたいと考えている。貴財団からの助成を賜り、心より感謝致します。

研究助成完了報告概要

(いずれも提出原文のまま、所属は提出時のもの)

■未就学聾者の実態調査とホームサインのアーカイブの作製

K28研XXI第494号

実施内容、成果および今後予想される効果の概要

日本手話の成立は、聾児が生まれた家庭で自然に手話が発生するというのが定説である。これをホームサインと呼ぶ。聾学校制度が確立した後も、経済的事情や地理的事情から聾学校に通わなかった聾者もいて、通称未就学聾者と呼ぶが、彼らはホームサインの使用者である。しかし標準手話を学んだ手話通訳者には理解できず、聾団体からも「手話がわからない人」「身振り」として処理されてきた。そのため、言語的特徴など、一切調査されていないのが現状である。

未就学聾者は離島に残存していることが言われてきたが、事前調査で新潟県佐渡地区に数名、在住していることがわかった。そこで筆者らは新潟市の聾団体や佐渡市の聾団体・通訳団体・聾家庭の協力を得て、まず実態調査を行い、家庭訪問などを実施し、ホームサインを収録・分析を行った。今回、2回目の収録を行った。

特徴的なホームサインを多く見つけることができた。基本的には、動作を伴う場合や、形状がある場合は、それ

木村 勉(豊田工業高等専門学校 情報工学科 教授)

を基準として表現している。これは主にジェスチャーであるが、手話の源流とも考えられる。

また、自分自身が中心である。例えば新潟に行くときには、海を渡って行くので、船の表現で、遠くの場所を示している。

日本手話との大きな違いの一つとして、色の表現が挙げられる。とにかく身近なモノを指さして代用する。例えば、「白」はティッシュペーパーを指差していた。しかし、常に表現したい色が身近にあるわけではない。日本手話では、必ずそこにある身体を使って表現している。「黒」は髪、「赤」は唇、「白」は歯といった表現であるが、これらは自分自身では見ることができないため、身近にあるモノを指さしていたと考えられる。

収録した動画は、ホームサインごとに分類し、保管している。今回収録に協力していただいた方は、すでに亡くなられてしまったので、貴重なアーカイブとして、後世に残すことができた。

■インフラ自動点検のためのマルチモーダルセンサを用いたオンライン診断

K29研XXII第499号

藤井 浩光(千葉工業大学 先進工学部 未来ロボティクス学科 准教授)

実施内容、成果および今後予想される効果の概要

道路トンネルをはじめとした社会インフラの老朽化は深刻な問題であり、適切な維持管理の実施が極めて重要な課題となっている。安全のためには、継続的なインフラの点検が必要不可欠であるが、点検対象の規模・件数はともに膨大な数に達しており、また点検員の人材が不足しているなどの背景から、維持管理を人手で行うことは困難となってきている。さらに、従来は看過されてきたような高所・狭所などの危険箇所を点検する必要が生じてきており、現在、ロボットなどを用いた自動点検システムの実現が強く望まれている。一般にこのようなプロセスを自動化するには、機械学習を用いることが多く、近年では膨大なデータを活用した深層学習により高度な識別問題の解決も自動化に成功している。しかし、深層学習などを用いて汎用的な問題を解決するには一般に膨大なデータが必要であるが、点検に関するデータベースの整備は現段階では進んでおらず、機械学習のための十分なデータを獲得することは困難な課題である。

人は初めて訪れる点検現場においても、信頼性の高い正確な診断を実施することが可能である。本研究で

は、この課題に対して熟練の点検員が行っている点検作業に注目し、人の診断プロセスのアルゴリズム化を行った。まず、従来のデータ駆動の学習アルゴリズムに対し、膨大なデータを必要としない教師なし・半教師あり学習を基にした診断アルゴリズムを提案した。具体的には、位置データや人手による予備知識を活用したクラスタリング手法を用いて、ラベル無しの少量のデータから健状と変状を識別する手法を提案した。さらに、視覚や聴覚情報以外の情報を用いたマルチモーダルな診断手法として、打音診断におけるハンマ反力の手応えなど触覚情報に相当する力センサのデータを活用するアプローチを提案し、アルゴリズム化した。

これらの学術的成果は、第36回日本ロボット学会学術講演会などで報告済みである。また、これらの研究活動はレスキュー分野などでも認められ、第14回競基弘賞学術業績賞を拝し、関連する研究活動に関する講演なども行った。今後、点検現場の自動化の必要性はますます高まっていくことが想定される。これらの研究を継続していくことで、現場において実用化可能な自動診断手法の構築を目指す。

■ 単元型を持つ概外延的な型理論の操作的意味論と型検査アルゴリズム

K29研XXII第501号

実施内容、成果および今後予想される効果の概要

CurienとDiCosmoが導入した書き換え系 $(\lambda\beta\eta\pi^*)'$ が、単元型を持つ概外延的な型理論の単純体系の操作的意味論を与えるには、 $(\lambda\beta\eta\pi^*)'$ の停止性が必要であるが、 $(\lambda\beta\eta\pi^*)'$ の停止性を証明するために、様々な技法を試み、学術雑誌に投稿し査読レポートを参考に、試行錯誤をした。

そのために、日本ソフトウェア科学界の第21回プログラミングおよびプログラミング言語ワークショップのプログラム委員や、コンピュータ科学と圏論の世話人を務め、TRSMeeetingなどで発表し、最近の研究の動向を調査した。本研究で検討した手法のいくつかを開発したBlanqui教授(仏INRIA)を2019年の12月18日から28日に招聘し、停止性の証明手法について説明を依頼し、講演会を開催し、 $(\lambda\beta\eta\pi^*)'$ の停止性の申請者の証明の吟味を依頼した。浜名誠氏(群馬大)、菊池健太郎氏

赤間 陽二(東北大学 大学院・理学研究科数学専攻 准教授)

(東北大)と議論した。

成果としては、 $(\lambda\beta\eta\pi^*)'$ の停止性の困難は(1)変数が書き換わり、(2)関数適用の形をした項がGirard-Taitの意味でcomputableならば、その頃の関数部分もcomputableになることであり、これを解決法がわかった。

BlanquiとDowekのグループは、可述的な依存積型理論による定理証明支援系に項書き換え系を支援系に導入し、その理論の型検査アルゴリズムとして、書き換え系の停止性をGeneralSchemaなどを用いた(自動的)証明と項の間の等式系の真偽判定を与えた。単位型をはじめ、概外延性を表す書き換え規則を、依存積型理論に導入した場合も、同様のアプローチが適用できることがわかった。以上に関する共著論文を現在準備している。単位型に限らず書き換え系を導入した依存積型理論の、型検査アルゴリズムが予想される。

■ モノラル受聴した音響信号からの音源方向・距離推定

K29研XXII第505号 鷗木 祐史(北陸先端科学技術大学院大学 先端科学技術研究科・情報科学系 教授)

実施内容

音響信号の音源到来方向・距離推定は、音源強調や音源分離、音響イベント検出に留まらず、ビッグデータにも利活用されている映像・音響監視システムといった様々な音響信号処理において重要なコア技術となっている。音源方向・距離推定にて、音空間情報(時間差やレベル差)を得るためには、複数のマイクロホン(マイクロホンアレイ)から得られる多チャンネル音響信号を利用しなければならない。しかし、マイクロホンの同一特性やAD変換の同期性といった条件をそろえなければ、その推定精度は著しく低下してしまう。

一方、ヒトは両耳で受聴した音から容易に音空間の情報を得ている。さらに片耳しか利用できないヒトでも、健聴者と同程度の音源方向・距離判断を行うことができる。申請者は、これまでに聴知覚特性の研究に取り組み、音の変調特性がこの能力に係わるものと予想し、その聴知覚メカニズムを一つ一つ検討してきた。その結果、音源信号と頭部伝達関数(頭部や耳介による回折特性)の変調特性ならびに頭部運動によるこれらの変分が音源方向推定に重要な役割を果たすことを明ら

かにした。本研究では、これらの結果に基づき、単一マイクロホンから得られたモノラル音響信号の変調成分からヒトと同じように音源方向・距離推定を可能にする方法を提案する。

成果

本研究では、まず東北大・RIECで提供されている頭部伝達関数(HRFT)データベース(約100名の両耳と音源との間の頭部インパルス応答)を利用し、音の方向推定に係わるHRFT上の変調伝達特性(HR-MTF)を分析した。その結果、水平角ならびに仰角の変化に関連したHR-MTF上の特徴を発見した。次に音源となる信号を音声と仮定し、その変調スペクトルの特徴を調べた。その結果、音声特有の特徴として変調周波数スペクトル上に有意なピーク(4 Hz付近)と傾斜がみられることを発見した。次に変調知覚の概念から音源の変調スペクトルとHR-MTFの線形システムとして受聴側の変調スペクトルを説明できると考え、音源方向推定(モノラルDOA)方法を検討した。ここでは音源の変調スペクトルの特徴(ピークと傾斜)ならびにHR-MTFの特性の関係

を関数モデルとして事前に学習し、その関数モデルから3次元の方向(水平角・仰角)を推定するモデルを構築した。未知データも含めたモノラル3D-DOAの推定精度を評価した結果、実効誤差で21.9度の精度をもつブラインド推定が可能であることがわかった。最後に、同様の方法でHR-MTFと受信した信号の音圧レベルの違いから、距離推定の実現可能性を検討した。事前に音源の絶対音圧レベルを定義できれば音源とマイク間の距離を推定可能であることがわかった。現在、推定精度を高くする方法を検討しているが、この点については今後の課題とした。本成果で得られたものは次の論文として掲載された。距離推定法の完成後、これらの成果をまとめたものをジャーナル論文として投稿予定である。

今後予想される効果の概要

提案法をベースに推定性が高精度であるモノラル音源方向・距離推定を実現できれば、映像・音響監視システムといった様々な応用技術に活用できるだけでなく、片耳を失った人への音響空間理解の機能補助といった補聴技術のイノベーションにもつながる。ヒトの聴力特性は左右耳で異なる。本研究の特色はアドホックマイクロフォンアレー処理で取り組む問題を、ヒトの聴覚メカニズムを理解することで解決するところにある。提案法により音源方向・距離推定を実現できれば、映像・音響監視システムといった様々な応用技術に活用できるだけでなく、片耳を失った人への音響空間理解の機能補助といった補聴技術のイノベーションにもつながる。

■視野領域の差異によるVR視聴時の生体影響

K29研XXII第507号

松浦 康之(岐阜市立女子短期大学 国際文化学科 専任講師)

実施内容、成果および今後予想される効果の概要

本研究ではHMDを用いたVR視聴において、視野領域のサイズの違いや視標の速度の違いが生体に与える影響について詳細に研究を行った。VR酔いの発生機序を解明することを目的に2種類の実験を行い、検討を行った。

実験1では、視野領域の違いについて検討を行った。重心動揺検査の結果から、視野狭窄がある映像視聴時の方が姿勢は安定し、動揺量が低下する結果が得られた。これは、視野狭窄により周辺視野の情報量が縮小されたことで、情報処理が容易になったことが考えられる。NIRS検査の結果から、視野狭窄の影響で酸素化ヘモグロビン濃度が前頭葉で有意に増加し、側頭葉の一部で有意に減少した。これらの結果から、視野狭窄に伴い腹側視覚路が賦活化して脳血流量が増加したと考えられる。また、前頭葉で脳血流量が増加したのは、視野狭窄によって体平衡を保持する上で参照している視覚情報量が減少して、重力方向を推量する上で補完すべき情報処理に伴い、小脳への負荷が過大になったと推察される。

実験2では、指標の速度の違いについて検討を行った。重心動揺検査の結果から、映像の違いによる差はみられなかった。このことから、視標の速度の違いは体平衡系に大きな影響を与えないことが示唆された。心電図検査の結果から、視標の速度の大きさが低い映像に比べ、視標の速度の大きさが高い映像視聴時に交感

神経が亢進した。これは、視標が速く運動することで視標の認識を困難にし、視標の把握のために集中力が増し、交感神経が賦活化した可能性が示唆される。NIRS検査の結果から、視標が速く運動することで視標の認識を困難にし、視標把握のために集中度が増し、前頭葉の脳血流量が増加した可能性が示唆された。また、後頭葉で脳血流量の減少がみられた。これは、指標の速度が上がった影響で、視覚対象物の精密な視覚処理が出来なかった可能性が考えられる。

以上の結果から、視野狭窄の影響あるいは、指標の速度の違いが、脳内の伝達経路あるいは、活性部位に変化が見られた。特に、視野狭窄時の映像視聴が生体に与える影響は殆ど検討されておらず、今回得られた結果は、今後研究を発展する上で、重要な方向性を得られた。

まとめ

本研究では、視野領域のサイズや視標の速度に着目し、VR視聴時の身体状態を、重心動揺検査、心電図検査、NIRS検査を用いて多角的に評価した。その結果、視野狭窄の影響あるいは、指標の速度の違いにより脳内の伝達経路あるいは、活性部位に変化などが見られた。特に、視野狭窄時の映像視聴が生体に与える影響は殆ど検討されておらず、今回得られた結果は、今後研究を発展する上で、重要な方向性を得ることが出来た。また、本研究の成果については、電気学会誌等への論

文投稿を予定している。

今後、今回の実験で得られたデータの更なる詳細な検討をしていくと同時に、映像提示時の背景や背景の

変化についても、検討を行う。これにより、3D映像視聴やVR視聴に関する新たな知見や視座を取得し、当該分野と産業の発展に貢献する。

■言語資源と画像資源を活用したゼロショットでの映像検索

K29研XXII第511号

実施内容、成果および今後予想される効果の概要

1. 実施内容、成果

本研究提案では、インターネット上に日々アップロードされる多種多様な大規模映像の中から、検索クエリに複数の概念を含んだ文(以下、クエリ文)を用いて詳細に映像を検索するため、以下の2つの技術的課題に取り組んだ。

研究項目1:クエリ文に合致した少量の映像からの潜在的概念の獲得

クエリ文に完全に合致した少量のデータ(画像・映像)を収集し、そのデータがすでにある概念とどの程度相関があるかを分析することにより、クエリ文中に直接に記述されていない潜在的概念を新たに獲得することを実験により確認した。例えば、クエリ文「destroyed buildings」を直接的に表現可能な概念識別器は存在しないという問題に対し、近い意味(言い換え)を表す「ruin([建物などの] 廃墟、遺跡)」が存在することが確認できた。また、画像の特徴としては近い(視覚的に近い)概念である「garbage heap(ゴミの山)」「dump(ゴミの山)」を用いることで、「destroyed buildings」という概念の検索精度を向上させることができることもわかった。

研究項目2:言語・画像・映像資源を用いた潜在的な概念の獲得

クエリ文から映像検索することを目的に、①「言語資

植木 一也(明星大学 情報学部情報学科 准教授)

源を用いた方法」、②「画像・映像資源を用いた方法」、③「双方を統合する方法」をそれぞれ検討し、映像検索に適した潜在的な概念を抽出する方式とそれを実現するために必要な資源の明確化に取り組んだ。

①については、Universal Sentence Encoderを活用し、クエリ文に近い概念を直接的に取得することを試みた。この手法により、検索キーワードと識別器名の対応付けを行う際、同じ識別器名を持つが意味が違った概念識別器を誤って選んでしまうケースを軽減できることがわかった。②と③は同時に取り組み、画像・映像と言語(クエリ文)を共通の一空間上に写像する手法を検討した。大規模な画像とキャプションデータベースを用いて学習を行い、画像と説明文を同一空間上に写像し、TRECVIDベンチマークの大規模映像を用いてクエリ文から直接的に映像検索を行えることも確認できた。

2. 今後予想される効果

ディープラーニング等の機械学習技術の発展により、学習済みの概念については高精度に検出できるが、新しいトレンドや新しい手口の犯罪等、未知の概念については正確に検出することができないという問題点がある。本研究の成果により、映像検索のクエリ文中には具体的に明示されていないが、人が知識として持っている潜在的概念を自動的に抽出することにより、新しい概念を検出可能とする仕組みが構築できることが期待される。

■深層学習に基づいた新しい信号処理技術の確立と歌声および楽器音生成への応用

K30研XXⅢ第531号

大浦 圭一郎(名古屋工業大学 大学院工学研究科 特任准教授)

(1)実施内容、成果および今後予想される効果の概要

従来のデジタル信号処理理論は、音声関連の研究分野では最も根本的な考え方として広く普及しているが、このような変換・処理で取り扱える枠組みの中に制限されていたため、モデル構造に関する過度の制約による性能限界があった。2016年、ニューラルネットワークを用いて音声波形を直接モデル化する手法としてWaveNetが提案されたが、過去の音声サンプル列から次の音声サンプルを生成する自己回帰構造を持つため、合成時に並列演算ができないことから、実時間で合成できない問題があった。また、補助特徴量として与えたピッチ情報とは異なるピッチの波形が生成されることがあった。2017年に提案されたParallel WaveNetでは、自己回帰構造を持たないため実時間で合成できるものの、指定のピッチを再現できない場合がある問題はまだまだ残っている。これらの問題に対し、深層学習に基づいた音声波形の直接モデル化手法を新たに確立するための研究の一つとして、明示的に周期信号と非周期信号の列を入力することで、対応する音声サンプルの列を一度に生成する手法を提案した。提案モデルを学習する際、画像変換の分野で高い性能を示しているCycleGANのサイクル構造を応用することで、音声波形と周期・非周期信号の相互変換の同時学習を行った。主観評価実験から、サイクル構造を持つ学習が合成歌声の品質を向上させることを示した。

本研究期間は、歌声および楽器音を対象として、従来のデジタル信号処理の枠を超えた音響モデリングの研究を進め、その一部を学術論文としてまとめ、投稿・発表することができた。その中でも、上記で説明した深層学習に基づいて周期・非周期信号からより自然な音声波形を生成する構成は、日本音響学会の粟屋潔学術奨励賞(2019年9月)や情報処理学会のマイクロソフト情報学研究賞(2020年3月)を受賞するなどの成果をあげている。また、それだけにとどまらず、近年のテキスト音声合成の分野でその有効性が示されている、入出力の時間的な対応情報をモデル自体が推定できるEnd-to-End構造を歌声の合成に応用する検討も進め、その一部を発表することができた。本研究により、これまでのデジタル信号処理における波形生成モデルとは全く異なる性質を持つ高性能なモデル化が期待でき、合成された歌声の品質も格段に向上することが期待される。今後も引き続き研究を進め、研究成果が出た段階で速

やかに学術論文としてまとめ、投稿・発表を行い、関連分野全体のレベルアップと当該技術の普及に寄与したい。

(2)実施内容および成果の説明

本報告書に同梱した4件の発表論文について、それぞれの内容と成果を示す。

【論文1】和田蒼汰, 法野行哉, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “歌声合成におけるニューラルボコーダの比較検討,” 音声研究会, vol. 119, no. 321, SP2019-42, pp. 85-90, 東京, 日本, 2019年12月.

本研究では、近年の音声合成分野において波形生成モジュールとして利用されるニューラルボコーダを歌声合成に用いた場合の性能を、主観評価実験および客観評価実験により比較した。過去に自身が出力した波形サンプル列を元に次の時刻の波形サンプルを予想する自己回帰型のニューラルボコーダである、WaveNetおよびWaveRNNの方式が自然性やピッチの再現性の面で他の手法より高い性能を出すことを確認した。

【論文2】Keiichiro Oura, Kazuhiro Nakamura, Kei Hashimoto, Yoshihiko Nankaku, and Keiichi Tokuda, “Deep neural network based real-time speech vocoder with periodic and aperiodic inputs,” 10th ISCA Speech Synthesis Workshop (SSW10), pp. 13-18, Vienne, Austria, September, 2019.

WaveNetなどに代表される近年のニューラルボコーダは、過去の波形サンプル列から次の波形サンプルを生成する自己回帰構造を持ち、合成時に並列演算できないことから、実時間で合成できない問題がある。また、学習する際のデータベースに無い音高の再現ができない問題や、指定したピッチ情報の音高を再現できるとは限らない問題があった。本研究では、明示的に周期信号と非周期信号の列を入力として用い、対応する波形サンプルの列を一度に生成する手法とその学習手法を提案し、歌声合成における主観評価実験によりその有効性を示した。

【論文3】大浦圭一郎, 高木信二, 中村和寛, 橋本佳, 南角吉彦, 徳田恵一, “周期・非周期信号を用いた敵対的生成ネットワークに基づくリアルタイム音声ボコーダ,”

日本音響学会2019年秋季研究発表会, pp. 997-998, 滋賀, 日本, 2019年9月.

論文2で提案したニューラルボコーダに、近年様々な分野で有効性が確認されている敵対的ネットワークを導入することで、深層学習の高度化を図った。

【論文4】村田舜馬, 藤本崇人, 法野行哉, 高木信二, 橋本佳, 大浦圭一郎, 南角吉彦, 徳田恵一, “楽譜時間情報を用いたアテンション機構に基づく歌声合成の検討,” 日本音響学会2019年秋季研究発表会, pp. 943-944, 滋賀, 日本, 2019年9月.

これまでの歌声合成では、楽譜から音響パラメータへの変換モジュール、発声タイミングの推定モジュール、お

よび波形生成モジュールなどが、それぞれ独立に定義されて最適化されてきた。より高品質な歌声を合成するためには、システム全体として最適化可能なEnd-to-Endモデルが必要と考えられる。音声合成の分野では近年、アテンション機構を含むsequence-to-sequenceモデルにより、音響パラメータと時間構造の同時モデリングが提案されているが、歌声合成のように一つのシーケンスが長い場合は、音響パラメータと時間構造の対応関係の探索空間が広がってしまい、そのまま流用することができない。本研究では楽譜の時間情報や音高情報を探索に利用することでこの問題を解決し、歌声合成においても音響パラメータと時間構造の同時モデリングが可能であることを示した。

■音源信号分離と立体音響にもとづいた選択的受聴に関する研究

K30研XXIII第534号

実施内容

本課題では、音源信号分離技術と立体音響再生技術を組み合わせた音響再生技術の拡張として、ユーザが聞きたい音やユーザに伝えたい音を選択的に提示する手法の確立に向けた検討を行った。実験では、聞きたい音を明らかとする前段階として、聞きたくない音、すなわち不快音に焦点を当て検討を進めた。研究期間内で実施した内容は、次のとおりである。

1) 不快音判定のための主観的検討

不快音であるか否かを決定するための基準を考えるため、候補となる音を定めて大規模アンケートを実施した。

2) 交通騒音の収録と分類

環境音の一つである道路交通環境音の収録、ならびに深層学習を用いた分類の検討を行った。

研究成果

不快音判定のための主観的検討では、20歳前後の若年層233名と、60歳以上のシニア層104名を被験者として、あらかじめ選定した不快音の候補を文字情報として表現し、この文字情報を対象に不快と感じられる割合

西野 隆典(名城大学 都市情報学部・都市情報学科 教授)

いについてアンケート調査を実施した。得られた成果として、身の回りに不快と感じられる音が存在すると回答した割合は2割程度であること、世代による差異がみられることが挙げられる。また、得られた成果は日本音響学会2020年春季研究発表会で報告した(*)。

交通騒音の収録と分類の検討では、名城大学ナゴヤドーム前キャンパス周辺の交通騒音の収録と、分類実験に向けた収録データの精査ならびに正解ラベルの付与を行った。また、得られたデータをもとに、深層学習を用いた分類実験の予備検討を行った。

(*)新型コロナウイルス感染症拡大防止対応のため発表会は中止となったが、講演論文集の発行をもってすべての発表が成立したものとみなされた。

今後予想される効果

不快と感じられる音について、実際に音を聞かせるのではなく、文字を用いて評価を行うことがある程度達成できたことから、例えばインターネット検索のキーワードのように、文字情報を使って不快な音を遮断することへの道筋をつけることができたと考えられる。

動 き

☆事務局日誌より☆

2019年

12.23

□K通信46号発送

3.5

□第20回評議員会成立(決議の省略)

2020年

2.5

□「ロボットシンポジウム2020名古屋」開催
吹上ホール

3.18

□第7回臨時理事会開催(WEB会議)

2.19

□第20回理事会成立(決議の省略)

3.20

□内閣府へ2020年度事業計画・
収支予算報告

CONTENTS

◇ 2020年度 助成金交付について	1
◇ 応募要領	2
◇ ロボットシンポジウム2020名古屋	3
◇ 会議開催報告	4
◇ フォーラム・シンポジウム等開催助成完了報告	5
◇ 研究助成完了報告概要	6